

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using ew model

Run on: January 30, 2004, 15:23:07 ; Search time 21 Seconds  
(without alignments)  
1124.260 Million cell updates/sec

Title: US-10-037-311B-1

Perfect score: 3004

Sequence: 1 MDQNSYRRSSPIRTTGGSS.....GTLVPHVHCEDISWGLKLV 558

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents, AA:\*

- 1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.psp:\*
- 2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.psp:\*
- 3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.psp:\*
- 4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.psp:\*
- 5: /cgn2\_6/ptodata/1/iaa/PTCUS\_COMB.psp:\*
- 6: /cgn2\_6/ptodata/1/iaa/backfiles1.psp:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description          |
|------------|-------|-------------|--------|----|----------------------|
| 1          | 111.5 | 3.7         | 804    | 4  | US-09-107-532A-6348  |
| 2          | 93    | 3.1         | 723    | 4  | US-09-134-001C-5060  |
| 3          | 92.5  | 3.1         | 468    | 4  | US-09-252-991A-31996 |
| 4          | 92.5  | 3.1         | 993    | 4  | US-08-836-687B-30    |
| 5          | 92    | 3.1         | 809    | 4  | US-09-186-276B-58    |
| 6          | 92    | 3.1         | 809    | 4  | US-08-842-445-58     |
| 7          | 92    | 3.1         | 809    | 4  | US-09-186-188B-58    |
| 8          | 91    | 3.0         | 554    | 4  | US-09-489-847-352    |
| 9          | 90    | 3.0         | 758    | 2  | US-08-222-617A-6     |
| 10         | 90    | 3.0         | 3666   | 2  | US-08-222-617A-12    |
| 11         | 90    | 3.0         | 3727   | 2  | US-08-222-617A-27    |
| 12         | 90    | 3.0         | 3778   | 2  | US-08-222-617A-2     |
| 13         | 89.5  | 3.0         | 1286   | 3  | US-09-268-140-3      |
| 14         | 89    | 3.0         | 1285   | 2  | US-08-540-406-6      |
| 15         | 89    | 3.0         | 1285   | 3  | US-08-656-055-6      |
| 16         | 89    | 3.0         | 1285   | 3  | US-08-954-668-6      |
| 17         | 89    | 3.0         | 1285   | 3  | US-08-918-658-6      |
| 18         | 89    | 3.0         | 1285   | 4  | US-09-724-631-6      |
| 19         | 89    | 3.0         | 1285   | 5  | PCT-US95-13233-6     |
| 20         | 88.5  | 2.9         | 515    | 3  | US-08-369-822C-23    |
| 21         | 88.5  | 2.9         | 515    | 3  | US-08-582-776C-38    |
| 22         | 88.5  | 2.9         | 515    | 3  | US-08-434-831B-35    |
| 23         | 88    | 2.9         | 957    | 4  | US-09-914-259-16     |
| 24         | 86.5  | 2.9         | 771    | 4  | US-09-462-284-2      |
| 25         | 86.5  | 2.9         | 1032   | 4  | US-09-914-259-26     |
| 26         | 85.5  | 2.8         | 1874   | 4  | US-09-331-403-2      |
| 27         | 85    | 2.8         | 461    | 4  | US-09-134-001C-5311  |

|    |    |     |      |   |                     |                    |
|----|----|-----|------|---|---------------------|--------------------|
| 28 | 85 | 2.8 | 584  | 1 | US-08-179-738-7     | Sequence 7, Appli  |
| 29 | 85 | 2.8 | 584  | 2 | US-08-628-145-7     | Sequence 7, Appli  |
| 30 | 85 | 2.8 | 591  | 1 | US-08-179-738-5     | Sequence 5, Appli  |
| 31 | 85 | 2.8 | 591  | 2 | US-08-628-145-5     | Sequence 5, Appli  |
| 32 | 85 | 2.8 | 596  | 1 | US-08-179-738-2     | Sequence 2, Appli  |
| 33 | 85 | 2.8 | 596  | 2 | US-08-628-145-2     | Sequence 2, Appli  |
| 34 | 85 | 2.8 | 742  | 4 | US-09-107-532A-4996 | Sequence 4996, Ap  |
| 35 | 85 | 2.8 | 777  | 2 | US-08-874-678-3     | Sequence 3, Appli  |
| 36 | 85 | 2.8 | 777  | 3 | US-08-643-839-3     | Sequence 3, Appli  |
| 37 | 85 | 2.8 | 777  | 4 | US-09-348-886-3     | Sequence 3, Appli  |
| 38 | 85 | 2.8 | 1298 | 1 | US-08-222-616-33    | Sequence 33, Appli |
| 39 | 85 | 2.8 | 1298 | 1 | US-08-340-011-2     | Sequence 2, Appli  |
| 40 | 85 | 2.8 | 1298 | 3 | US-08-901-710-2     | Sequence 33, Appli |
| 41 | 85 | 2.8 | 1298 | 4 | US-08-446-648-33    | Sequence 33, Appli |
| 42 | 85 | 2.8 | 1298 | 5 | PCT-US95-04228-33   | Sequence 33, Appli |
| 43 | 85 | 2.8 | 1363 | 1 | US-08-340-011-4     | Sequence 4, Appli  |
| 44 | 85 | 2.8 | 1363 | 2 | US-08-874-678-32    | Sequence 32, Appli |
| 45 | 85 | 2.8 | 1363 | 3 | US-08-643-839-32    | Sequence 32, Appli |

#### ALIGNMENTS

#### RESULT 1

US-09-107-532A-6348

; Sequence 6348, Application US/09107532A

; Patent NO. 6583275

; GENERAL INFORMATION:

; APPLICANT: Lynn A Doucette-Stamm and David Bush

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS

; NUMBER OF SEQUENCES: 7310

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: GENOME THERAPEUTICS CORPORATION

; STREET: 100 Beaver Street

; CITY: Waltham

; STATE: Massachusetts

; COUNTRY: USA

; ZIP: 02354

; COMPUTER READABLE FORM:

; MEDIUM TYPE: CD/ROM ISO9660

; COMPUTER: PC

; OPERATING SYSTEM: <unknown>

; SOFTWARE: ASCII

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/107,532A

; FILING DATE: 30-Jun-1998

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 60/085,598

; FILING DATE: 14 May 1998

; APPLICATION NUMBER: 60/051571

; FILING DATE: July 2, 1997

; ATTORNEY/AGENT INFORMATION:

; NAME: Ariniello, Pamela Deneke

; REGISTRATION NUMBER: 40,489

; REFERENCE/DOCKET NUMBER: GTC-012

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (781)893-5007

; TELEFAX: (781)893-8277

; INFORMATION FOR SEQ ID NO: 6348:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 804 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; HYPOTHETICAL: YES

; ORIGINAL SOURCE:

; ORGANISM: Enterococcus faecium

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: (B) LOCATION 1...804

; SEQUENCE DESCRIPTION: SEQ ID NO: 6348:

US-09-107-532A-6348

Query Match 3.7%; Score 111.5; DB 4; Length 804;  
Best Local Similarity 21.4%; Pred. No. 0.019;  
Matches 101; Conservative 61; Mismatches 132; Indels 179; Gaps 26;

QY 101 LIASGFDEOCLSRYSOVHYRK-----PSYPKSSYLISKLNRYEKLHKRCOPGT-----150  
DB 194 LMAKFPHD-----EYRKQRELAVAAYK-----TIRTKQLFVEAPTGTGKTI 237  
QY 151 ESYKKAALKQDQDHIDGDGE-----CKVVMWISPSGLGNRLSLASVFLYALLTD 200  
DB 238 STLFPALKAVGEE-----EGEKIFYLTAKTITRQVAEDAMTALXDTGAEVKSVTLTA--KD 291  
QY 201 RVLLVDRGKMDLDFCEPFLGMSWLLPLDPFMTDQ-----FDGLNQEISSR 245  
DB 292 KICF-----LDTECNP-----DQCPYANGYNNRINEGLWDLNHNEN-- 328  
QY 246 CYGYMVKNQVIDT---EGTLSHLYLHLVHDYGDHDMFCEGDDQTFIGKVPMLIVKTDNY 302  
DB 329 ----QITREVIETVARKHTLCPPELSL-----DVSIMCD-----VIIGDYNV 366  
QY 303 -FVPSLWLPFGFDELNKLFPQKATVPHHLGRVLPHTNQVWGLVTR---YIEAYLSHAD 358  
DB 367 LFDPTVYLRRFFDEKNE-----DYLFL-LIDEAHLNVRNSREMYSAELSSEYK 412  
QY 359 EKIGIQVRVDFDEPGFQHVMDQISSCTQKEKLLPEVDTL--VERSRHVNTPKHKAVALVT 416  
DB 413 TK-----RSKEAIPKEFKLHRFN-----KLLKEFDSIRETAKEDHWDYHOKAPAES 461  
QY 417 SLNAGY--AENLKSMTWEYPTSTGEIIGVHQPQSE-----GYQQ 453  
DB 462 LVKAGYQLSEKIKEWLAEPF-----EHPQEQLLPYVDFLLHFLKVPSEYVDHYET 512  
QY 454 TEKMHNGKALAE-----MYLLSLTDN-----LVTSAMSTFGYVAGGLGLK 495  
DB 513 TVEKTVDLIVKEFCIDPSLFLQSLDGKSGSLFSASFPLSYOETLGGQK 565

RESULT 2  
US-09-134-001C-5060  
; Sequence 5060, Application US/09134001C  
; Patent No. 6380370  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
; FILE REFERENCE: GTC-007  
; CURRENT APPLICATION NUMBER: US/09/134,001C  
; CURRENT FILING DATE: 1998-08-13  
; PRIOR FILING DATE: 1997-11-08  
; PRIOR APPLICATION NUMBER: US 60/064,964  
; PRIOR FILING DATE: 1997-11-08  
; PRIOR APPLICATION NUMBER: US 60/055,779  
; NUMBER OF SEQ ID NOS: 5674  
; SEQ ID NO 5060  
; LENGTH: 723  
; TYPE: PRT  
; ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-5060

Query Match 3.1%; Score 93; DB 4; Length 723;  
Best Local Similarity 20.9%; Pred. No. 1.4;  
Matches 110; Conservative 60; Mismatches 168; Indels 188; Gaps 28;

QY 113 SRYQSVHYRKPSYPKSSYLIS-----KLNRYEKLHRCRGGTESYKALKQLDQE--HID 166  
DB 212 SRQDEINYPKPKYKYLINSVDGYDLKL-NQKRYKD-----KKELELIEHIEKHQOE 262  
QY 167 GDCECKYVVMWISPSGLGNRLSLASVFLYALLTDRLVLDVRGKMDLDFCEPFLGMSWLL 226  
DB 263 G-----KILEV-----KGNNK-----KSYAQ 278  
QY 227 PLDFPMTDQDGLNQESSRCYGMVKNQVIDTBTGLSHLY-LHLVHDYGDHDKMFFCEGD 285

DB 279 PL-FNLTD-----LQOEAYKRYKMGPK-----ETLNTLQHLRYERHKLVTYPRTDSNYLTDDM 329  
QY 286 QTFIGKVPMLIVKTDNYFVPSLWLPFGFDELNKLFPQKATVF-----HH-----330  
DB 330 VDTIQLERAILATD-----YKSHVRDLISKFSFSSKGMHIFNNQKVSDDHAIIPTEVR 381  
QY 331 -----LGRYLFHPTNQVWGLVTRV-YEAYLSHAD-----EKIGIQVRVDE 370  
DB 382 PSTEQLSQREFKIYMLIAERFLENLMNPVLYEVLTHAQKDYHFLVKEKIPQL-----436  
QY 371 DPGPFQHVMDQISSCTQ-----KEKLLPEVDTLVERSRHVNTPKH--KAVLVTSL-----418  
DB 437 ---GYKALKDQLSSHTLTHSFKEGQLFKVHRIEIHETKAPVFNESGLLKAMENPQNH 493  
QY 419 ----NAGYAEHLKSMWEYPTSTGEIIGV-----HQPSEGYQOTTEKQWNGKA 463  
DB 494 IDLNDKXYAKTLK-----HTGGIGTATRADIIIEKLFNMNALESRDGKIKVTSKGKQ 545  
QY 464 LAEMYLSSLTDNLVTSAM-STFGYVAGGLGGLKPMI-----LYRPE 503  
DB 546 ILELSSELTSPILTAQWEEKMLIEKGKYNKSFQIEMKNFTFKVNVNKKSSQKXKHD 605  
QY 504 NRTTPD-PSCGRAM-----SMEPCFHSPPFPYDCCKATGIDTGT 540  
DB 606 NLTTTECPTCGKPMIKVKTNGQMLVC-QDP-----CKTKKNIQKRT 647

RESULT 3  
US-09-252-991A-31996  
; Sequence 31996, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; CURRENT APPLICATION NUMBER: 107196.136  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 31996  
; LENGTH: 468  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-31996

Query Match 3.1%; Score 92.5; DB 4; Length 468;  
Best Local Similarity 23.9%; Pred. No. 0.8;  
Matches 37; Conservative 26; Mismatches 57; Indels 35; Gaps 7;

QY 137 RNYEKLHRCRGGTESYKALKQLDQEHIDGDCECKYVVMWISFSGLNRLSLASVFLYA 196  
DB 110 RNYDRILE--GPIKTAISKAKQLKQSRLEHGTTSIIWF-----VNNGYTALDHDALLT 163  
QY 197 LITDRV-----LLVD-----RGKMDLDFCEPFLGMSWLLPLDPFMTDQDGL-----239  
DB 164 LIAHRVRNDTNEIDGIIVSGCYFHSDFSDFLWPFYVPIINDLKOPF-----EFDDLHRAW 220  
QY 240 NQESSRCYGM-----VKNQVIDTBTGLSHL 265  
DB 221 NLATRSMTALMQKPFQKQEIKGVPVDTQDIDNI 255

RESULT 4  
US-08-836-687B-30  
; Sequence 30, Application US/08836687B  
; Patent No. 6448034  
; GENERAL INFORMATION:  
; APPLICANT: Gasson, Michael John

```
; APPLICANT: Dodd, Helen Mair
; TITLE OF INVENTION: PRODUCTION OF VARIANT NISIN
; FILE REFERENCE: 20747/70
; CURRENT APPLICATION NUMBER: US/08/836,697B
; CURRENT FILING DATE: 1998-11-20
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 30
; LENGTH: 993
; TYPE: PRT
; ORGANISM: Lactococcus sp.
; US-08-836-687B-30

Query Match      3.1%; Score 92.5; DB 4; Length 993;
Best Local Similarity 19.1%; Pred. No. 2.7; Indels 167; Gaps 23;
Matches 90; Conservative 63; Mismatches 152;

Qy 89 VTNINSDKLLGGLASGDEDCSLRYQSVHVRKP--SPYKSSYLLSKLRNYEKHKRC 146
Db 224 ISNLOKD-----LLSDFSWNTFLTKVEAIDDKYIIPKKVKQFIOEYSEIE-----I 272
Qy 147 GPTESYKALKOLOQ-----EHIDGGECKVWVWISFSGLGNRLSLAS 191
Db 273 GEGIEKLKEIYQMSQILENDNYIQIDLISDSEINFDVKQKQOOLHAEFLGNTTKSVRR 332
Qy 192 VFL--YALLTDRVLL---VDRGKQMDLFCPEFLGMSWLLPLDFPMTQDFG-----LN 240
Db 333 TVLDDY---KDFIEKYGVQEQVITELPDSFP-GIGAPYNNHPRNDFYESEPTLYS 388
Qy 241 QESSRCYGYM---VKN-QVIDTEGLSH-----LYLHVHDYGDHDKM 279
Db 389 EEEREKYSMYEAVGNHNVINLDDLESHYQKMDLEKKSELQGLFLNLAKY---EKD 445
Qy 280 FCEGQDTGKVPMLIVKTDNYFVPSLWLIQPDDELNKLFPQKATVPHLGRYLFRT 339
Db 446 IFILGD-----IVGNNN-----LGGASGRFSALSP----- 471
Qy 340 NQVWGLVTRYEAYLSHADEKIGIQVRVDEDPGFQHVMDQISSCTQEKLLPVDTLV 399
Db 472 -----LTSYHRTIVDSVER-----ENENKEITSC-----EIVFLP 501
Qy 400 ERSRHYN-----TPKHKAVLVTSLNAGYAENLKSIMYWEYPTSTGEIIGV 443
Db 502 ENIRHANVMTSIRMRKVLPPFTSTSHNEVLLTNIVIGIDE--KEKFYARDISTQEVLEKF 559
Qy 444 HPSQEGYQOTEKQMN-----GKALAEMLLSLTNVLNVTSAWS-----TFGYV 487
Db 560 YITS-----MYNKTLSNELRFLYEISLDDKFGNLPWELIYRDPDYI 601

RESULT 5
US-09-186-276B-58
; Sequence 58, Application US/09186276B
; Patent No. 6388173
; GENERAL INFORMATION:
; APPLICANT: Benfey, Philip
; APPLICANT: DiLaurenzio, Laura
; APPLICANT: Wysocka-Diller, Joanna
; APPLICANT: Malamy, Jocelyn E.
; APPLICANT: Fysh, Leonard
; APPLICANT: Helariutta, Yrjo
; TITLE OF INVENTION: Scarecrow Gene, Promoter and Uses Thereof
; FILE REFERENCE: 5914-075-999
; CURRENT APPLICATION NUMBER: US/09/186,276B
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: 08/842,445
; PRIOR FILING DATE: 1997-04-24
; PRIOR APPLICATION NUMBER: 08/638,617
; PRIOR FILING DATE: 1996-04-26
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 58
; LENGTH: 809
; TYPE: PRT
; ORGANISM: Plant
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(809)
; OTHER INFORMATION: Xaa = Any Amino Acid
; US-08-842-445-58

Query Match      3.1%; Score 92; DB 4; Length 809;
Best Local Similarity 21.7%; Pred. No. 2.2; Indels 179; Gaps 18;
Matches 84; Conservative 52; Mismatches 179;

Qy 66 PSDSNRIMGFAEARVLD-----AGVFP-----NVTNINSDKLLGGLASGDE---- 108
Db 63 PPSSNQTGLANGFYLDLDFSSLDPPPEAYPSQNNNNNNINNKAVAGDLSSSDDADFS 122
Qy 109 DSCLSRQSVHYRKPSPYKSSYLLS-KLRNYEK-LHKRCGPGTESYKALKQLDQEH- 165
Db 123 DSVLKYSIQVLMEEEDMEKPCMFHDALALQAAEKSLYEALGKDPSSSSASSVDHPERLA 182
Qy 166 --DGDGCKYVWVWISFSGLGNRLSLASVFLYALLTORVLLVDRGKQMDLFCPEFLGMS 223
Db 183 SHSPDGSCS-----CGAFSDYAS-----TTTTSSDSHWSVDGLENRP-----S 221
Qy 224 WL---LPLDFPMTQDFGLNQESSRCYGYMYKNQVIDTEGLTSLHYLHLVHDYGDHDKMF 280
Db 222 WLHTPMSNF---VFQTSRSNSVTGGGGGNSAVYSGFGDDLVSNNPKD--DELAHQ 275
Qy 281 FCEGQD---TPGKVPMLIVKTDNYFVPSLWLIQPDDELNKLFPQKATVPHLGRY 334
Db 276 PKGVEEASKPLPKSSQLFIDVDSY-IP--MNSGSKENGSEVFVKTEKDETEHHHHS 331
Qy 335 LFHPTNQVGLVTRYEAYLSHADEKIGIQVRVDEDPGFQHVMDQISSCTQEKLLPE 394
Db 332 YAPPNNRLTGKSHRDEDEDFVEERSNKOSAVTYER-SELSEMFDMNFCGPGKVC-- 388
Qy 395 VDTLVERSRYNTPKHKAVLVTSLNAG 421
Db 389 -----ILNQNFPTESAKVVTQAQSNQ 408

RESULT 6
US-08-842-445-58
; Sequence 58, Application US/08842445A
; Patent No. 6441270
; GENERAL INFORMATION:
; APPLICANT: Benfey et al.
; TITLE OF INVENTION: Scarecrow Gene, Promoter and Uses
; FILE REFERENCE: 5914-056-999
; CURRENT APPLICATION NUMBER: US/08/842,445A
; CURRENT FILING DATE: 1997-04-24
; EARLIER APPLICATION NUMBER: 08/638,617
; EARLIER FILING DATE: 1996-04-26
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 58
; LENGTH: 809
; TYPE: PRT
; ORGANISM: Plant
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(809)
; OTHER INFORMATION: Xaa = Any Amino Acid
; US-08-842-445-58

Query Match      3.1%; Score 92; DB 4; Length 809;
Best Local Similarity 21.7%; Pred. No. 2.2; Indels 179; Gaps 18;
Matches 84; Conservative 52; Mismatches 179;

Qy 66 PSDSNRIMGFAEARVLD-----AGVFP-----NVTNINSDKLLGGLASGDE---- 108
Db 63 PPSSNQTGLANGFYLDLDFSSLDPPPEAYPSQNNNNNNINNKAVAGDLSSSDDADFS 122
```

QY 109 DSCLSRYSQSVHYRKPSPYKPSYLLIS-KLRNVEK-LHKRCGPGCTESYKALKQLDOEHI- 165  
DB 123 DSVLKYSIQLVMEEDMEERPCMFHDALALQAAEKSYEALGEKDPSSSASSVDHFERIA 182  
QY 166 --DGDGECKYVVMVIFSPGLNRIILSLASVFLYALLTDRLVLLVDRGKMDLDFCEPFLGMS 223  
DB 183 SHSPDGSCS-----GGAFSDYAS-----TTTTSSDSHWSVDGLENRP-----S 221  
QY 224 WL---LPLDPMTDQDGLNQESSRCYGYWVKQVQVDTGTGLSHLYLHLVHDYGDHDKMF 280  
DB 222 WLHTPMPSNF-----VFQSTSRNSVTVGGGGGNSAVYGSFGDGLVSNMFKD--DELAHQ 275  
QY 281 FCEGDO---TFIGKVPWLIVKTDNYFVPSLWLIIPGFDDELNKLKLF---POKATVFHILGRY 334  
DB 276 FKGVVEASKFLPKSQSLFIDVDSY-IP---MNSGSKENGSEVVFVTEKDETEHHHHS 331  
QY 335 LFHPTNQVGLVTRYEAYLSHADEKIGIQVRVDEDPGPFQHVMDQISSCCTQKEKLLPE 394  
DB 332 YAPPPNRLTGKXSHRDEDEDFVEERSNKQSAVYVEE-SELSEMFDMFLCGPKPVC--- 388  
QY 395 VDTLVERSHTVTPKHKAVLVTSNAG 421  
DB 389 -----ILNQNPFTESAKVVTQAQNSG 408

## RESULT 7

US-09-186-1888-58  
; Sequence 58, Application US/091861888  
; Patent No. 6455672  
; GENERAL INFORMATION:  
; APPLICANT: Benfey et al.  
; TITLE OF INVENTION: Scarecrow Gene, Promoter and Uses  
; FILE REFERENCE: 5914-074-999  
; CURRENT APPLICATION NUMBER: US/09/186,188B  
; PRIOR FILING DATE: 1998-11-05  
; PRIOR APPLICATION NUMBER: 08/842,445  
; PRIOR FILING DATE: 1997-04-24  
; PRIOR APPLICATION NUMBER: 08/638,617  
; NUMBER OF SEQ ID NOS: 79  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 58  
; LENGTH: 809  
; TYPE: PRT  
; ORGANISM: Plant  
; NAME/KEY: VARIANT  
; LOCATION: (1)...(809)  
; OTHER INFORMATION: Xaa = Any Amino Acid  
US-09-186-1888-58

Query Match 3.1%; Score 92; DB 4; Length 809;  
Best Local Similarity 21.7%; Pred. No. 2.2;  
Matches 84; Conservative 52; Mismatches 179; Indels 72; Gaps 18;  
QY 66 PSDSRIMGFAARVLD-----AGVFP-----NVTWINSKLLGLLASGFDE----- 108  
DB 63 PFSSNQTLGANGFYLDLDFSLDPEAYPSQNNNNNNNNKAVAGDLSSSSDADFS 122  
QY 109 DSCLSRYSQSVHYRKPSPYKPSYLLIS-KLRNVEK-LHKRCGPGCTESYKALKQLDOEHI- 165  
DB 123 DSVLKYSIQLVMEEDMEERPCMFHDALALQAAEKSYEALGEKDPSSSASSVDHFERIA 182  
QY 166 --DGDGECKYVVMVIFSPGLNRIILSLASVFLYALLTDRLVLLVDRGKMDLDFCEPFLGMS 223  
DB 183 SHSPDGSCS-----GGAFSDYAS-----TTTTSSDSHWSVDGLENRP-----S 221  
QY 224 WL---LPLDPMTDQDGLNQESSRCYGYWVKQVQVDTGTGLSHLYLHLVHDYGDHDKMF 280  
DB 222 WLHTPMPSNF-----VFQSTSRNSVTVGGGGGNSAVYGSFGDGLVSNMFKD--DELAHQ 275

QY 281 FCEGDO---TFIGKVPWLIVKTDNYFVPSLWLIIPGFDDELNKLKLF---POKATVFHILGRY 334  
DB 276 FKGVVEASKFLPKSQSLFIDVDSY-IP---MNSGSKENGSEVVFVTEKDETEHHHHS 331  
QY 335 LFHPTNQVGLVTRYEAYLSHADEKIGIQVRVDEDPGPFQHVMDQISSCCTQKEKLLPE 394  
DB 332 YAPPPNRLTGKXSHRDEDEDFVEERSNKQSAVYVEE-SELSEMFDMFLCGPKPVC--- 388  
QY 395 VDTLVERSHTVTPKHKAVLVTSNAG 421  
DB 389 -----ILNQNPFTESAKVVTQAQNSG 408  
RESULT 8  
US-09-489-847-352  
; Sequence 352, Application US/09489847  
; Patent No. 6476195  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: 98 Human Secreted Proteins  
; FILE REFERENCE: P2031P1  
; CURRENT APPLICATION NUMBER: US/09/489,847  
; EARLIER FILING DATE: 2000-01-24  
; EARLIER APPLICATION NUMBER: PCT/US99/17130  
; EARLIER FILING DATE: 1999-07-29  
; EARLIER APPLICATION NUMBER: 60/094,657  
; EARLIER FILING DATE: 1998-07-30  
; EARLIER APPLICATION NUMBER: 60/095,486  
; EARLIER FILING DATE: 1998-08-05  
; EARLIER APPLICATION NUMBER: 60/096,319  
; EARLIER FILING DATE: 1998-08-12  
; EARLIER APPLICATION NUMBER: 60/095,454  
; EARLIER FILING DATE: 1998-08-06  
; EARLIER APPLICATION NUMBER: 60/095,455  
; EARLIER FILING DATE: 1998-08-06  
; NUMBER OF SEQ ID NOS: 376  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 352  
; LENGTH: 554  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; NAME/KEY: SITE  
; LOCATION: (16)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (109)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-489-847-352

Query Match 3.0%; Score 91; DB 4; Length 554;  
Best Local Similarity 19.6%; Pred. No. 1.5;  
Matches 86; Conservative 50; Mismatches 152; Indels 150; Gaps 19;  
QY 119 HYRKPSYK-----PSSYLISKLRNVEKLRKRCGCT-ESYKALKQLDOEHIDGDECK 172  
DB 157 HWEESRSRKESRGSPRRGVALLRP-EPLHR-----GTADTLNVRVKL-----PCQ 202  
QY 173 YVWVWISFSGLGNRIILSLASVFLYALLTDRLVLLVDRGKMDLDFCEPFLGMSWLLPLDFPM 232  
DB 203 ITSYLVAHTLGRRLYPGVSVLLQKALMPVLLQGOAR-----L 240  
QY 233 TQOFDGLNQESSRCYGYWVKQVQVDTGTGLSHLYLHLVHDYGDHDKMFCEGQTFIGKV 292  
DB 241 VEEENGRRAKLACDGNEDITMTFVDRRGTAEPQGOQLV-----ICCEGNAGFY-EV 290  
QY 293 PMLIVKTDNYFVPSLWLIIPGFDDELNKLKFPQKAT-----VFHILGRYLFHPTNOV- 342  
DB 291 GCUSTPLEAGYSVLGNHFGFAGSTGVPFPQNEANAMDVVQFAIHLG---FQPDIII 347  
QY 343 -----W-GLVTRYEAY--LSH 356

Db 348 YANSIGFTATWAMSGYPOVSAMILDASDDLVPLALKVMPSWRGLVTRTRVQHNLN 407  
QY 357 ADEKIGIQVRVDFEDPQ---FQWMDQISSCTQKEKLPE--VDTLVRSRH---VNT 407  
Db 408 AEQLCRYQ-----GPVLLIRTKDEIITTVPEDIMSRGNDLLKLLQHRYPVWA 459  
QY 408 PKKAVLVTSLNAGYAENLUSMY--WE-----YPTSTGEIIGVHQ 445  
Db 460 EGLRVVRQWLEASSQLEASYSRWEVEDMCLSVLRSYQAEHGPDPFWSVGEDMSADG 519  
QY 446 PSQEGYQOTEKKHNGKA 463  
Db 520 RROLALFLARKHLHNEA 537

## RESULT 9

US-08-222-617A-6  
; Sequence 6, Application US/08222617A  
; Patent No. 5882879  
; GENERAL INFORMATION:  
; APPLICANT: Veenstra, Annemarie E.  
; APPLICANT: Martin, Juan F.  
; APPLICANT: Garcia, Bruno D.  
; APPLICANT: Gutierrez, Santiago  
; APPLICANT: Barredo, Jose L.  
; APPLICANT: Von Doehren, Hans  
; APPLICANT: Palissa, Harriet  
; APPLICANT: Van Liempt, Henk  
; APPLICANT: Montenegro, Eduardo P.  
; TITLE OF INVENTION: A Method for Influencing Beta-Lactam  
; TITLE OF INVENTION: Antibiotic Production and for Isolation of Large  
; TITLE OF INVENTION: Quantities of ACV Synthetase  
; NUMBER OF SEQUENCES: 27  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff  
; STREET: 300 South Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/222,617A  
; FILING DATE: 04-APR-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; REFERENCE/DOCKET NUMBER: 97,157  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 758 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FRAGMENT TYPE: internal  
; ORGANISM SOURCE:  
; ORGANISM: Penicillium chrysogenum  
; FEATURE:  
; NAME/KEY: Domain  
; LOCATION: 1..758  
; OTHER INFORMATION: /label= Domain II  
; OTHER INFORMATION: /note= "Domain II of ACV Synthetase from  
; OTHER INFORMATION: Penicillium chrysogenum; aa 1397-2154"  
US-08-222-617A-6

Query Match 3.0%; Score 90; DB 2; Length 758;

Best Local Similarity 20.7%; Pred. No. 3.2;

Matches 67; Conservative 35; Mismatches 99; Indels 122; Gaps 13;

QY 71 RIMGFAEAVLDAGVFPNVTNINSKLLGLGSLGASGDESDCLSRYSQVHYRKPSPKPS- 129  
Db 132 RIKGWAASGTL--LYPSVLPANPD-----SKWS---VSNPSPLSRST 168  
QY 130 --SYLISKRYEKLHKRCGPGTSYKALKQDOEHIDGGECKYVWISFSGLCNRIL 187  
Db 169 DLAYIITYT-----SGTGRPKG-----VTVEHHGVVNLQV 198  
QY 188 SLASVFLYALLTRVLLVDRGKMD---DLFCEPFLGMSWLLPLDFFPMTDQDGLNQESS 244  
Db 199 SLKVFGLRDTDDDEVILSFSNVFDFHFEQMTDAILNGQTLVLN-----DGRGDK 251  
QY 245 RYGYMVKNQVIDTGTSLHLYLHLVHDYGDHDKMFFCEGD-----QTFIGKVP 293  
Db 252 RLYRYIEKNRVYLSGTPSVVSMYEFSEFKDHLRRVDCVGEAFSPVFDKIRETFHGLV- 310  
QY 294 WLIKTDNYFVPSLWLIPIGFDDELNKLPPQKATVPHLGRYLPHFTNQWGLVTRYEAY 353  
Db 311 -----INGYGTVEVSITTHKRLYPF-PERRM----- 335  
QY 354 LSHADEKIGIOVR-----VFDED 371  
Db 336 -----DKSIGQQVHNSTSYVLNED 354

## RESULT 10

US-08-222-617A-12  
; Sequence 12, Application US/08222617A  
; Patent No. 5882879  
; GENERAL INFORMATION:  
; APPLICANT: Veenstra, Annemarie E.  
; APPLICANT: Martin, Juan F.  
; APPLICANT: Garcia, Bruno D.  
; APPLICANT: Gutierrez, Santiago  
; APPLICANT: Barredo, Jose L.  
; APPLICANT: Von Doehren, Hans  
; APPLICANT: Palissa, Harriet  
; APPLICANT: Van Liempt, Henk  
; APPLICANT: Montenegro, Eduardo P.  
; TITLE OF INVENTION: A Method for Influencing Beta-Lactam  
; TITLE OF INVENTION: Antibiotic Production and for Isolation of Large  
; TITLE OF INVENTION: Quantities of ACV Synthetase  
; NUMBER OF SEQUENCES: 27  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff  
; STREET: 300 South Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/222,617A  
; FILING DATE: 04-APR-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; REFERENCE/DOCKET NUMBER: 97,157  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3666 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Penicillium chrysogenum  
; FEATURE:

; NAME/KEY: Protein  
; LOCATION: 1..3666  
; OTHER INFORMATION: /label= region  
; OTHER INFORMATION: /note= "Region of ACV Synthetase; aa 62-3727"  
US-08-222-617A-12

Query Match 3.0%; Score 90; DB 2; Length 3666;  
Best Local Similarity 20.7%; Pred. No. 43;  
Matches 67; Conservative 35; Mismatches 99; Indels 122; Gaps 13;

QY 71 RINGFAEAVLDAGVPNPVNTINSQKLLGGLLASGDEDSCLSRYSQSVHYRKSPYKPS- 129  
DB 1467 RIKGMAASGTL---LYPSVL PANPD-----SKWS---VSNESPLSRST 1503  
QY 130 --SYLISKLRNVEKLHRCGPGTESYKALKQLDQEHIDGDECKYVWMISFSGLGNRIL 187  
DB 1504 DLAYIIYT-----SGTGRPKG-----VTVEHGVNVLQV 1533  
QY 188 SLASFLYALLTRVLLVDRGKMD---DLFCBPFLGMSWLLPLDPFMTDQFGLNQESS 244  
DB 1534 SLSKVFLGRDDEVLISFNSYVDFHFVEQMTDAILNGQTLVLN-----DGMRGDKE 1586  
QY 245 RCYGVNKNQVIDTEGTLSHLYLHLDYDHDHMFCEGD-----QTFEGKVP 293  
DB 1587 RLRYIEKRVLYSGTSPSVVSMYBFSRFDKHLRRVDCVGEAFSEPVDFDKIRETFHGLV- 1645  
QY 294 WLIVKTDNYFVPSLWLPDGLNKLFPKATVFFHILGRYLFHPTNQVWGLVTRYVEAY 353  
DB 1646 -----INGYGPTEVSITTHKRLYPF-PERRM----- 1670  
QY 354 LSHADEKIGIQVR-----VPDED 371  
DB 1671 -----DKSIGQOVHNSSTSYVLNED 1689

RESULT 11  
US-08-222-617A-27  
; Sequence 27, Application US/08222617A  
; Patent No. 5882879  
; GENERAL INFORMATION:  
; APPLICANT: Veenstra, Annemarie E.  
; APPLICANT: Martin, Juan F.  
; APPLICANT: Garcia, Bruno D.  
; APPLICANT: Gutierrez, Santiago  
; APPLICANT: Barredo, Jose L.  
; APPLICANT: Von Doehren, Hans  
; APPLICANT: Palissa, Harriet  
; APPLICANT: Van Liempt, Henk  
; APPLICANT: Montenegro, Eduardo P.  
; TITLE OF INVENTION: A Method for Influencing Beta-Lactam  
; TITLE OF INVENTION: Antibiotic Production and for Isolation of Large  
; TITLE OF INVENTION: Quantities of ACV Synthetase  
; NUMBER OF SEQUENCES: 27  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff  
; STREET: 300 South Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/222,617A  
; FILING DATE: 04-APR-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; REFERENCE/DOCKET NUMBER: 97,157  
; INFORMATION FOR SEQ ID NO: 27:  
; SEQUENCE CHARACTERISTICS:

; LENGTH: 3727 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-222-617A-27

Query Match 3.0%; Score 90; DB 2; Length 3727;  
Best Local Similarity 20.7%; Pred. No. 44;  
Matches 67; Conservative 35; Mismatches 99; Indels 122; Gaps 13;

QY 71 RINGFAEAVLDAGVPNPVNTINSQKLLGGLLASGDEDSCLSRYSQSVHYRKSPYKPS- 129  
DB 1528 RIKGMAASGTL---LYPSVL PANPD-----SKWS---VSNESPLSRST 1564  
QY 130 --SYLISKLRNVEKLHRCGPGTESYKALKQLDQEHIDGDECKYVWMISFSGLGNRIL 187  
DB 1565 DLAYIIYT-----SGTGRPKG-----VTVEHGVNVLQV 1594  
QY 188 SLASFLYALLTRVLLVDRGKMD---DLFCBPFLGMSWLLPLDPFMTDQFGLNQESS 244  
DB 1595 SLSKVFLGRDDEVLISFNSYVDFHFVEQMTDAILNGQTLVLN-----DGMRGDKE 1647  
QY 245 RCYGVNKNQVIDTEGTLSHLYLHLDYDHDHMFCEGD-----QTFEGKVP 293  
DB 1648 RLRYIEKRVLYSGTSPSVVSMYBFSRFDKHLRRVDCVGEAFSEPVDFDKIRETFHGLV- 1706  
QY 294 WLIVKTDNYFVPSLWLPDGLNKLFPKATVFFHILGRYLFHPTNQVWGLVTRYVEAY 353  
DB 1707 -----INGYGPTEVSITTHKRLYPF-PERRM----- 1731  
QY 354 LSHADEKIGIQVR-----VPDED 371  
DB 1732 -----DKSIGQOVHNSSTSYVLNED 1750

RESULT 12  
US-08-222-617A-2  
; Sequence 2, Application US/08222617A  
; Patent No. 5882879  
; GENERAL INFORMATION:  
; APPLICANT: Veenstra, Annemarie E.  
; APPLICANT: Martin, Juan F.  
; APPLICANT: Garcia, Bruno D.  
; APPLICANT: Gutierrez, Santiago  
; APPLICANT: Barredo, Jose L.  
; APPLICANT: Von Doehren, Hans  
; APPLICANT: Palissa, Harriet  
; APPLICANT: Van Liempt, Henk  
; APPLICANT: Montenegro, Eduardo P.  
; TITLE OF INVENTION: A Method for Influencing Beta-Lactam  
; TITLE OF INVENTION: Antibiotic Production and for Isolation of Large  
; TITLE OF INVENTION: Quantities of ACV Synthetase  
; NUMBER OF SEQUENCES: 27  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff  
; STREET: 300 South Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/222,617A  
; FILING DATE: 04-APR-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; REFERENCE/DOCKET NUMBER: 97,157  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:

```
/ ; LENGTH: 3778 amino acids
/ ; TYPE: amino acid
/ ; TOPOLOGY: linear
/ ; MOLECULE TYPE: protein
US-08-222-617A-2

Query Match 3.0%; Score 90; DB 2; Length 3778;
Best Local Similarity 20.7%; Pred. No. 45;
Matches 67; Conservative 35; Mismatches 99; Indels 122; Gaps 13;

QY 71 RINGFAEAVILAGVPPNTNINSKLLGLLASGDEDSCLSRYSQSVHYRKPSPYKPS- 129
Db 1528 RIKGMAAGTL---LYPSVLPANPD-----SKMS---VSNPSPLSRST 1564
QY 130 --SYLISKLRYNEKLHKRCGPGETSYKALKQDQEHIDGCECKVYVWISFGIGNRL 187
Db 1565 DLAYIIYT-----SGTGPCKG-----VTVEHGVVNVNQV 1594
QY 188 SLASVFLYALLDRVLLVDRGKMD---DLFCEPFLGMSWLLPLDPFMTDQDGLNQESS 244
Db 1595 SLKVFGLRDTDDDEVILSFNTYVDFHFVEQMTDAILNGQTLVLN-----DGMRGDKE 1647
QY 245 RCYGVYKNOVIDTGTLSHLVHLVHDYGDHDKMFFCEGD-----QTFIGKVP 293
Db 1648 RLRYIEKRVLYLSTGTSVSMYEFSPKDLHRRVDCVGEAFSEPVFDKIRETFHGLV- 1706
QY 294 WLIVKTDNYFVPSLWLIQGFDDDELNKLFPQKATVPHLGRYLFHPTNQVWGLVTRYEAY 353
Db 1707 -----INGYGTVEVSIITHKRLYFF-PEREM----- 1731
QY 354 LSHADEKIGIQVR-----VFDEB 371
Db 1732 -----DKSIGQQVHNSTSYLNEB 1750

RESULT 13
US-09-268-140-3
; Sequence 3, Application US/09268140
; Patent No. 6268176
; GENERAL INFORMATION:
; APPLICANT: Gemmill, Robert M.
; APPLICANT: Drabkin, Harry A.
; TITLE OF INVENTION: TRC8, A GENE RELATED TO THE HEDGEHOG RECEPTOR, PATCHED
; FILE REFERENCE: 93445-00004
; CURRENT APPLICATION NUMBER: US/09/268.140
; CURRENT FILING DATE: 2000-03-12
; PRIOR APPLICATION NUMBER: US 60/077,723
; PRIOR FILING DATE: 1998-03-12
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1286
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-268-140-3

Query Match 3.0%; Score 89.5; DB 3; Length 1286;
Best Local Similarity 20.6%; Pred. No. 8.7;
Matches 64; Conservative 41; Mismatches 81; Indels 125; Gaps 17;

QY 218 PFLGMSW-----LLPLDFPMTDQDGLN-----QESSRCYGYM 250
Db 672 PFLMRSWKFLTVMGFLAALISLYASTRLQDGLDIIDLVPKDSNEHKFLDAQTRLFQFY 731
QY 251 VKNOVIDTEGTLSH-LYLHLVHDYGDHDKMFFCEGDQTFIGKVPWLVKTDNYFVPSLWL 309
Db 732 SMYAV--TQGNFEYPTQQLRDY--HD-----SFV-RVPH-VIKNDNGGLPDFWL 776
QY 310 IPGFDELNKLFPQKATVPHLGRYLFHPTNQVWGLVTRYEAYLSHADEKIGIQVRVFD 369
Db 777 L-----LF-----SEWLNQKIFD 791
QY 370 EDPGFQHVMDQISSCTQKEKLLPEVD-----LVERSRHVNTPKHAVLTVS--LNA 420

Query Match 3.0%; Score 89; DB 2; Length 1285;
Best Local Similarity 20.3%; Pred. No. 9.8;
Matches 63; Conservative 39; Mismatches 83; Indels 126; Gaps 16;

QY 218 PFLGMSW-----LLPLDFPMTDQDGLN-----QESSRCYGYM 250
Db 672 PFLMRSWKFLTVMGFLAALISLYASTRLQDGLDIIDLVPKDSNEHKFLDAQTRLFQFY 731
QY 251 VKNOVIDTEGTLSH-LYLHLVHDYGDHDKMFFCEGDQTFIGKVPWLVKTDNYFVPSLWL 309
Db 732 SMYAV--TQGNFEYPTQQLRDY--HDSF-----RVPH-VIKNDNGGLPDFWL 775
QY 310 IPGFDELNKLFPQKATVPHLGRYLFHPTNQVWGLVTRYEAYLSHADEKIGIQVRVFD 369
Db 776 L-----LF-----SEWLNQKIFD 790
QY 370 EDPGFQHVMDQISSCTQKEKLLPEVD-----LVERSRHVNTPKHAVLTVS--LNA 420
Db 791 E-----EYRDGLTKECFFNASSDAILAYKLIVQTGHVDNPDVKELVLTNRLVNS 841
```





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2004, 15:34:17 ; Search time 39 Seconds

(without alignments)  
2973.483 Million cell updates/sec

Title: US-10-037-311B-1

Perfect score: 3004

Sequence: 1 MDQSYRRSSPIRTTGG.....GTLVPHVRHCDISWGLKLV 558

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 789580 seqs, 207824079 residues

Total number of hits satisfying chosen parameters: 789580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description         |
|------------|-------|-------------|--------|-------|---------------------|
| 1          | 1492  | 49.7        | 533    | 12    | US-10-338-777-35    |
| 2          | 994.5 | 33.1        | 595    | 12    | US-10-359-194A-152  |
| 3          | 110   | 3.7         | 658    | 12    | US-09-882-227-612   |
| 4          | 103.5 | 3.4         | 409    | 12    | US-10-238-075-835   |
| 5          | 99    | 3.3         | 769    | 11    | US-09-934-455-504   |
| 6          | 96    | 3.2         | 520    | 12    | US-10-032-201B-305  |
| 7          | 96    | 3.2         | 524    | 12    | US-10-032-201B-303  |
| 8          | 96    | 3.2         | 528    | 12    | US-10-032-201B-304  |
| 9          | 96    | 3.2         | 934    | 12    | US-10-369-493-19663 |
| 10         | 93    | 3.1         | 1418   | 12    | US-10-369-493-19663 |
| 11         | 92    | 3.1         | 809    | 15    | US-10-253-007-58    |
| 12         | 92    | 3.1         | 946    | 12    | US-10-369-493-19347 |
| 13         | 91    | 3.0         | 536    | 12    | US-10-369-493-646   |
| 14         | 90.5  | 3.0         | 419    | 10    | US-09-843-905A-6    |
| 15         | 90.5  | 3.0         | 419    | 12    | US-10-317-250-6     |

|    |      |     |      |    |                     |                    |
|----|------|-----|------|----|---------------------|--------------------|
| 16 | 90.5 | 3.0 | 419  | 15 | US-10-197-666A-134  | Sequence 134, Appl |
| 17 | 90.5 | 3.0 | 502  | 9  | US-09-815-242-13660 | Sequence 13660, A  |
| 18 | 90.5 | 3.0 | 512  | 12 | US-10-369-493-10428 | Sequence 10428, A  |
| 19 | 90.5 | 3.0 | 847  | 12 | US-10-369-493-6478  | Sequence 6478, Ap  |
| 20 | 90.5 | 3.0 | 903  | 12 | US-10-317-832-11    | Sequence 11, Appl  |
| 21 | 90.5 | 3.0 | 903  | 12 | US-10-104-047-2460  | Sequence 2460, Ap  |
| 22 | 90   | 3.0 | 634  | 12 | US-10-032-585-7524  | Sequence 7524, Ap  |
| 23 | 89.5 | 3.0 | 1286 | 10 | US-09-898-533-3     | Sequence 3, Appli  |
| 24 | 89.5 | 3.0 | 1331 | 12 | US-10-369-493-2641  | Sequence 2641, Ap  |
| 25 | 89   | 3.0 | 756  | 12 | US-10-369-493-19945 | Sequence 19945, A  |
| 26 | 89   | 3.0 | 1285 | 8  | US-08-954-701A-6    | Sequence 6, Appli  |
| 27 | 89   | 3.0 | 1285 | 12 | US-09-754-032-6     | Sequence 6, Appli  |
| 28 | 89   | 3.0 | 1285 | 12 | US-10-421-446-6     | Sequence 6, Appli  |
| 29 | 88.5 | 2.9 | 420  | 14 | US-10-041-030-4     | Sequence 4, Appli  |
| 30 | 88.5 | 2.9 | 420  | 15 | US-10-197-666A-136  | Sequence 136, App  |
| 31 | 88.5 | 2.9 | 522  | 15 | US-10-128-714-3043  | Sequence 3043, Ap  |
| 32 | 88.5 | 2.9 | 1693 | 12 | US-10-120-801-107   | Sequence 107, App  |
| 33 | 88   | 2.9 | 957  | 12 | US-10-080-608A-16   | Sequence 16, Appl  |
| 34 | 88   | 2.9 | 957  | 12 | US-10-370-685-105   | Sequence 105, App  |
| 35 | 87.5 | 2.9 | 748  | 12 | US-10-369-493-6312  | Sequence 6312, Ap  |
| 36 | 87.5 | 2.9 | 911  | 12 | US-10-032-585-7655  | Sequence 6, Appli  |
| 37 | 87   | 2.9 | 417  | 12 | US-09-793-705-6     | Sequence 6, Appli  |
| 38 | 87   | 2.9 | 418  | 10 | US-09-843-905A-4    | Sequence 4, Appli  |
| 39 | 87   | 2.9 | 418  | 12 | US-10-317-250-4     | Sequence 4, Appli  |
| 40 | 87   | 2.9 | 418  | 12 | US-10-443-108-10    | Sequence 10, Appl  |
| 41 | 87   | 2.9 | 418  | 14 | US-10-041-030-2     | Sequence 2, Appli  |
| 42 | 87   | 2.9 | 516  | 12 | US-10-032-585-7407  | Sequence 7407, Ap  |
| 43 | 87   | 2.9 | 522  | 15 | US-10-128-714-8043  | Sequence 8043, Ap  |
| 44 | 86.5 | 2.9 | 875  | 12 | US-10-369-493-10263 | Sequence 10263, A  |
| 45 | 86.5 | 2.9 | 876  | 9  | US-09-815-242-12623 | Sequence 12623, A  |

## ALIGNMENTS

## RESULT 1

US-10-338-777-35  
; Sequence 35. Application US/10338777  
; Publication No. US20030188343A1  
; GENERAL INFORMATION:  
; APPLICANT: Lynx Therapeutics, Inc.  
; APPLICANT: United States Department of Agriculture  
; APPLICANT: Bowen, Benjamin A  
; APPLICANT: Haudenschild, Christian D  
; APPLICANT: Buckler, Edward S  
; TITLE OF INVENTION: Identification of Genes Associated with Growth in Plants  
; FILE REFERENCE: 37-000510US  
; CURRENT APPLICATION NUMBER: US/10/338,777  
; CURRENT FILING DATE: 2003-01-07  
; NUMBER OF SEQ ID NOS: 405  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 35  
; LENGTH: 533  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
US-10-338-777-35

Query Match 49.7%; Score 1492; DB 12; Length 533;  
Best Local Similarity 54.6%; Pred. No. 7.6e-147;  
Matches 295; Conservative 90; Mismatches 133; Indels 22; Gaps 9;  
QY 29 LQMKYLSGTMKLTFTTCLIVFSLVAFSMIFQHPDSNRIMGFA---EARVLDAVG 85  
Db 1 MQKFOISG-KIVKTLGKMKVLIAVSGSLF-----ILSYNNFNKKLDATT 49  
QY 86 FPNV--TNINSKLLGLASGFDESDCLSRYSQSVH-YRKPSYPKPSYLSKLRNVEKL 142  
Db 50 KVDIKETERPKVGLIGLLTADFEGSLSRVHKYFLYKPKPFPSEVLVSKLSYEML 109  
QY 143 HKRCGPGTESYKALKQLDQEH-DCDGECKYVWISFSGLNRLISLASVFLYALLTDR 201  
Db 110 HKRCGDPTEYKEAEIKLSRDDASBSNGECRYVWVAGYGLGNRLTLTSLASVFLYALLTER 169



|    |     |  |     |
|----|-----|--|-----|
| Qy | 206 | DRGKMDDDLFCBFFFLGMSWLLPLDPFMTDQF----                         | 236 |
| Db | 176 | TEVEKQEDFYNOYWRAMEERPQNETLNNRFRVHVLTIKIGIPNEKRVVEAFKYRQKK    | 235 |
| Qy | 237 | -----DGLNOESRCYGVY-----KNQVIDTEGTLSH-----                    | 276 |
| Db | 236 | GIEIEDLLKDLQKYCYFCQIAFKBEDDKLNFALSFLVLEMDVIYPLLELAYSQDK-     | 294 |
| Qy | 277 | DKMFFCEGDQTEIGKVPWLIVKTDNYFVPSLWLIPG----                     | 322 |
| Db | 295 | -----GVLKQD--FPIIYLYESYICRRVCGIGTSLNKLFP--                   | 335 |
| Qy | 323 | --OKATVFHHLGRYLF-----HPTNOVMGLV-----                         | 362 |
| Db | 336 | HIQOBYEYFKSLKAHFVCLUTERQRPNNDEFKKLFITIDPVYKKNKYFLERLENFDTKEP | 395 |
| Qy | 363 | IQVRVFDEDPGGFQHVMDQISSCQKEKLLPEVDLVERSRYV--NTPKHKAVLVTSNA    | 420 |
| Db | 396 | V-----DTQKCNIEHIMPQTLPBQRDLGSGFOAHEKYLHTGN-----LTLTGYS       | 444 |
| Qy | 421 | GYAEN-----LKSM-----  | 437 |
| Db | 445 | KYSNNSFQEKRDMEKFGKQSSILKNQSLKOLESGEKEIEKRASDLADWALKIWTYPILE  | 504 |
| Qy | 438 | GIILGVHQPQSBGYOQTEBKMMHNGKALAEMYLLS                          | 471 |
| Db | 505 | AETLBEYKPKPEKKEKEEYKLBKGGKVVYDLS                             | 538 |

## RESULT 4

```

RESULTS 4
US-10-238-075-835
; Sequence 835, Application US/10238075
; Publication No. US20030148324A1
; GENERAL INFORMATION:
; APPLICANT: I.N.S.E.R.M.
; TITLE OF INVENTION: Polynucleotides which are of nature B2/D+ A- and which are isolated
; TITLE OF INVENTION: E.coli, and biological uses of these polynucleotides and of their
; FILE REFERENCE: BLANDINE
; CURRENT APPLICATION NUMBER: US/10/238,075
; CURRENT FILING DATE: 2002-09-10
; PRIOR APPLICATION NUMBER: 0003145
; PRIOR FILING DATE: 2000-03-10
; NUMBER OF SEQ ID NOS: 1576
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 835
; LENGTH: 409
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-238-075-835

```

## RESULT 5

```

US-09-934-455-504
; Sequence 504, Application US/09934455
; Publication No. US20030121070A1
; GENERAL INFORMATION:
; APPLICANT: Adam, Luc
; APPLICANT: Creelman, Robert
; APPLICANT: Dubell, Arnold
; APPLICANT: Heard, Jacqueline
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Keddle, James
; APPLICANT: Pilgrim, Marsha
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Reuber, Lynne
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Pineda, Omar
; TITLE OF INVENTION: Genes for Modifying Plant Traits IV
; FILE REFERENCE: MBI-0025
; CURRENT APPLICATION NUMBER: US/09/934,455
; CURRENT FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/227439
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: MBI-0022
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: MBI-0023
; PRIOR FILING DATE: 2001-04-17
; NUMBER OF SEQ ID NOS: 516
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 504
; LENGTH: 769
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-934-455-504

```

## RESULT 6

RESULT 6  
US-10-032-201B-305  
; Sequence 305, Application US/10032201B  
; Publication No. US20030167524A1  
; GENERAL INFORMATION:  
; APPLICANT: Van Rooijen, Gijb

```

; APPLICANT: Deckers, Harm
; APPLICANT: Heifetz, Peter Bernard
; APPLICANT: Briggs, Steven
; APPLICANT: Dalmia, Bipin Kumar
; APPLICANT: Del Val, Greg
; APPLICANT: Zaplachinski, Steve
; APPLICANT: Moloney, Maurice
; TITLE OF INVENTION: METHODS FOR THE PRODUCTION OF MULTIMERIC PROTEINS, AND RELATED
; FILE REFERENCE: 38814 351B
; CURRENT APPLICATION NUMBER: US/10/032,201B
; CURRENT FILING DATE: 2001-12-19
; NUMBER OF SEQ ID NOS: 313
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 305
; LENGTH: 520
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-032-201B-305

Query Match      3.2%; Score 96; DB 12; Length 520;
Best Local Similarity 21.2%; Pred. No. 1.3;
Matches 58; Conservative 40; Mismatches 99; Indels 76; Gaps 12;

QY 282 CEGDQTFIGKVPMLIVKTDNYFVPSLWLIQFDDELNKLFPQ-----KATVPFHL 331
DB 229 CAGFLTGIGLDTVMRMS-----IP-----LRGFDQMSSLVTEHMGHTQFLKGCVPVSHI 280
QY 332 GRYLFHPTNQVWGLVTRYEAYLSHADEKIGIQVRVDEDPGPPQHVMQDISSCTQKEKL 391
DB 281 KKL---PTNQL-----QVTWEDHASK-----EDTGTFTVLWAIGR----- 314
QY 392 LPEVDTL-VERSRHVTPKHAKVLVTSLNAGYAENLKSMTWEYPTSTGEIIGVHQPQSEG 450
DB 315 VPETRTLNLKAGISTPNKQKIIVDAQEA-----TSVPHIYAIGDVAEGR 360
QY 451 YQOTEKMHNGKALAE-MYLLSLT-----DNLVTSAMSTFGYVAQGLGGLKPWIL----- 499
DB 361 PELTPTAIKAGKLLAQRLFGKSSLTMDYSNVPTTFTFPLEYGCGLSEEEAVALHGQEHV 420
QY 500 -----YRPNRTTPDPSCGRAMSMEPCFHSPP 526
DB 421 EVYHAYYKPLEFTVADRDASQCYIKMCMREPP 453

RESULT 7
US-10-032-201B-303
; Sequence 303, Application US/10032201B
; Publication No. US20030167524A1
; GENERAL INFORMATION:
; APPLICANT: Van Rooijen, Gijb
; APPLICANT: Deckers, Harm
; APPLICANT: Heifetz, Peter Bernard
; APPLICANT: Briggs, Steven
; APPLICANT: Dalmia, Bipin Kumar
; APPLICANT: Del Val, Greg
; APPLICANT: Zaplachinski, Steve
; APPLICANT: Moloney, Maurice
; TITLE OF INVENTION: METHODS FOR THE PRODUCTION OF MULTIMERIC PROTEINS, AND RELATED
; FILE REFERENCE: 38814 351B
; CURRENT APPLICATION NUMBER: US/10/032,201B
; CURRENT FILING DATE: 2001-12-19
; NUMBER OF SEQ ID NOS: 313
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 303
; LENGTH: 524
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 527
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-032-201B-304

Query Match      3.2%; Score 96; DB 12; Length 528;
Best Local Similarity 21.2%; Pred. No. 1.3;
Matches 58; Conservative 40; Mismatches 99; Indels 76; Gaps 12;

QY 282 CEGDQTFIGKVPMLIVKTDNYFVPSLWLIQFDDELNKLFPQ-----KATVPFHL 331
DB 237 CAGFLTGIGLDTVMRMS-----IP-----LRGFDQMSSLVTEHMGHTQFLKGCVPVSHI 288
QY 332 GRYLFHPTNQVWGLVTRYEAYLSHADEKIGIQVRVDEDPGPPQHVMQDISSCTQKEKL 391
DB 289 KKL---PTNQL-----QVTWEDHASK-----EDTGTFTVLWAIGR----- 322
QY 392 LPEVDTL-VERSRHVTPKHAKVLVTSLNAGYAENLKSMTWEYPTSTGEIIGVHQPQSEG 450
DB 323 VPETRTLNLKAGISTPNKQKIIVDAQEA-----TSVPHIYAIGDVAEGR 368
```

```

US-10-032-201B-303

Query Match      3.2%; Score 96; DB 12; Length 524;
Best Local Similarity 21.2%; Pred. No. 1.3;
Matches 58; Conservative 40; Mismatches 99; Indels 76; Gaps 12;

QY 282 CEGDQTFIGKVPMLIVKTDNYFVPSLWLIQFDDELNKLFPQ-----KATVPFHL 331
DB 233 CAGFLTGIGLDTVMRMS-----IP-----LRGFDQMSSLVTEHMGHTQFLKGCVPVSHI 284
QY 332 GRYLFHPTNQVWGLVTRYEAYLSHADEKIGIQVRVDEDPGPPQHVMQDISSCTQKEKL 391
DB 285 KKL---PTNQL-----QVTWEDHASK-----EDTGTFTVLWAIGR----- 318
QY 392 LPEVDTL-VERSRHVTPKHAKVLVTSLNAGYAENLKSMTWEYPTSTGEIIGVHQPQSEG 450
DB 319 VPETRTLNLKAGISTPNKQKIIVDAQEA-----TSVPHIYAIGDVAEGR 364
QY 451 YQOTEKMHNGKALAE-MYLLSLT-----DNLVTSAMSTFGYVAQGLGGLKPWIL----- 499
DB 365 PELTPTAIKAGKLLAQRLFGKSSLTMDYSNVPTTFTFPLEYGCGLSEEEAVALHGQEHV 424
QY 500 -----YRPNRTTPDPSCGRAMSMEPCFHSPP 526
DB 425 EVYHAYYKPLEFTVADRDASQCYIKMCMREPP 457

RESULT 8
US-10-032-201B-304
; Sequence 304, Application US/10032201B
; Publication No. US20030167524A1
; GENERAL INFORMATION:
; APPLICANT: Van Rooijen, Gijb
; APPLICANT: Deckers, Harm
; APPLICANT: Heifetz, Peter Bernard
; APPLICANT: Briggs, Steven
; APPLICANT: Dalmia, Bipin Kumar
; APPLICANT: Del Val, Greg
; APPLICANT: Zaplachinski, Steve
; APPLICANT: Moloney, Maurice
; TITLE OF INVENTION: METHODS FOR THE PRODUCTION OF MULTIMERIC PROTEINS, AND RELATED
; FILE REFERENCE: 38814 351B
; CURRENT APPLICATION NUMBER: US/10/032,201B
; CURRENT FILING DATE: 2001-12-19
; NUMBER OF SEQ ID NOS: 313
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 304
; LENGTH: 528
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 527
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-032-201B-304

Query Match      3.2%; Score 96; DB 12; Length 528;
Best Local Similarity 21.2%; Pred. No. 1.3;
Matches 58; Conservative 40; Mismatches 99; Indels 76; Gaps 12;

QY 282 CEGDQTFIGKVPMLIVKTDNYFVPSLWLIQFDDELNKLFPQ-----KATVPFHL 331
DB 237 CAGFLTGIGLDTVMRMS-----IP-----LRGFDQMSSLVTEHMGHTQFLKGCVPVSHI 288
QY 332 GRYLFHPTNQVWGLVTRYEAYLSHADEKIGIQVRVDEDPGPPQHVMQDISSCTQKEKL 391
DB 289 KKL---PTNQL-----QVTWEDHASK-----EDTGTFTVLWAIGR----- 322
QY 392 LPEVDTL-VERSRHVTPKHAKVLVTSLNAGYAENLKSMTWEYPTSTGEIIGVHQPQSEG 450
DB 323 VPETRTLNLKAGISTPNKQKIIVDAQEA-----TSVPHIYAIGDVAEGR 368
```



```

; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 5914-074-999
; CURRENT APPLICATION NUMBER: US/10/253.007
; CURRENT FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: US/09/186,188
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: 08/842,445
; PRIOR FILING DATE: 1997-04-24
; PRIOR APPLICATION NUMBER: 08/638,617
; PRIOR FILING DATE: 1996-04-26
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 58
; LENGTH: 809
; TYPE: PRT
; ORGANISM: Plant
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(809)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-253-007-58

Query Match      3.1%; Score 92; DB 15; Length 809;
Best Local Similarity 21.7%; Pred. No. 6.9;
Matches 84; Conservative 52; Mismatches 179; Indels 72; Gaps 18;

Qy 66 PSDSNRTMGFAEARVLD-----AGVRP-----NVTNINSDKLLGLLGLASGDFE--- 108
Db 63 PPSNQITGLANGFYLDLDFSSLDPEAYPSQNNNNNNNNKAVAGDLLSSSSDDADFS 122
Qy 109 DSCLSRYQSVHYRKPSYKPSYLLS- KLRNYEK- LHKRCGPGTGYESYKALKQJDOBH1- 165
Db 123 DSVLYKISQVLMDEEMKPCMFHDALQAAEKLGEALGKDPSSSSASSVDHPERLA 182
Qy 166 --DGGCKYVWVWISFSGLNRIILSLASVFLYALLTORVLLVDRKMDLDFCEPFLGMS 223
Db 183 SHSPDGSCS-----GGAFSDYAS-----TTTTSSDSHWSVDGLNRP-----S 221
Qy 224 WL---LPLDRPMTQDFGLNQESSRCYGMVKNQVIDTEGTLHLHLVHDYGDHDKMF 280
Db 222 WLHTPMENP-----VFQSTSRNSVUTGGCGNSAVTSGSGDGLVSNMFKD--DELAQ 275
Qy 281 FCEGDQ---TPIGKVPWLIVKTDNYFVPSLWLIPIFGDDELNKLFP---PQKATVPHHLGRY 334
Db 276 FKKGVEEASKFLPKSSQLFIDVDSY-IP---MNSGSKENGSEVFPVTEKKDTEHHHHS 331
Qy 335 LFHPTNQVGLVTRYEAYLSHADEKIGIQVRVDEDPGFQHVMDQISSCTQKELLP 394
Db 332 YAPPNNLTGKSHWRDEDEDFVEERSNKQSAVYVEE-SELSEMFNDNMLFLCGPGKPYC-- 388
Qy 395 VDTLVERSRRVHTPKHKAIVLTSINAG 421
Db 389 -----ILNQNFTESAKVVTQASNG 408

RESULT 12
US-10-369-493-19347
; Sequence 19347, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374

Query Match      3.0%; Score 91; DB 12; Length 536;
Best Local Similarity 19.8%; Pred. No. 4.4;
Matches 105; Conservative 63; Mismatches 166; Indels 150; Gaps 26;

; SEQ ID NO 19347
; LENGTH: 946
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-10-369-493-19347

Query Match      3.1%; Score 92; DB 12; Length 946;
Best Local Similarity 21.7%; Pred. No. 8.9;
Matches 105; Conservative 63; Mismatches 166; Indels 150; Gaps 26;

Qy 20 SKSVNFSQLQMKYLSGTMKLTFTTTCILVSVLVAFSMIFQHPHSDS-NRIMGAEAA 78
Db 376 AETNLISGL--KGYTTGGTV-----HVINNVQVGTFTDHDSSRSLYSATA 420
Qy 79 RVLDAGVFPNVNINSD-----KLLGLLASGDFE-----SCLSYQSVHYRKPS 124
Db 421 QMLDIPVF-----HVGDDPEACVHIAKLVAEYQT-FKSDVIDLVYRRYGHNEGDEPS 475
Qy 125 PYKPSSYLI-----SKLRNYEKLHKRCGPGTESYKALKOLDQEHID-----G 167
Db 476 FTQPMYVDIIRKHTPTVRTLYAAKLAQNKI-----PAEES--EAIKQRCQOEFDALTRA 528
Qy 168 DGECKYVWVWISFSGL-----GNRILSLASVFLYALLTORVLLVDR-----GK 209
Db 529 RQESQPKPSALEGLMKPYQGGALKSAPDV---STAVDKQVLCDAKRLSTLPEGFNVHR 585
Qy 210 DMDLDFCEPFLGM-----SWLLPLDPFMTDQDFGLNQESSRCYGMVKNQVIDTE-GTL 262
Db 586 DVERTVKKRLGMLDSGELQW-----SEGSLAYATLLSEGYNIRITGQDCERGTP 636
Qy 263 SHLYLHLVHDYGDHDKMFPCEGQDTFGKVPWLIVKTDNYFVPSLWLIPIFGDDELNKLFP 322
Db 637 SHRHA-VVHD-----VKTGEKFPVLRQFISG----- 661
Qy 323 QKATVPHHLGRYLFH-----PTNQVGLVTRYEAYLSHADEKIGIQVRVDEDPGFQHV 378
Db 662 -----TGRNGFHIYNSPLSEM-GVLGFEGYSLDVPDGLTAAQAFQGFDRNGA-QII 711
Qy 379 MDQ- ISSCTQKELLPEDVTLVERSRRVHTPKH-KAVLVTSLNAGYAENLKSMTWYPTPS 436
Db 712 IDQFIAGESKWRRLSGLTLLPHGYEGQGFPEHSSARLERFLDLCAEDNIQVC---YPTT 768
Qy 437 TGEI 440
Db 769 PAQI 772

RESULT 13
US-10-369-493-646
; Sequence 646, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 646
; LENGTH: 536
; TYPE: PRT
; ORGANISM: Deinococcus radiodurans
US-10-369-493-646

Query Match      3.0%; Score 91; DB 12; Length 536;
Best Local Similarity 19.8%; Pred. No. 4.4;
Matches 105; Conservative 63; Mismatches 166; Indels 150; Gaps 26;
```



